INFECTION AND IMMUNITY, Dec. 1993, p. 5147–5156
0019-9567/93/125147-10$02.00/0

Vol. 61, No. 12

Characterization of *Clostridium perfringens* Iota-Toxin Genes and Expression in *Escherichia coli*

SYLVIE PERELLE, MARYSE GIBERT, PATRICE BOQUET, AND MICHEL R. POPOFF*

Laboratoire des Toxines Microbiennes, Institut Pasteur, 28 Rue du Dr. Roux, 75724 Paris Cedex 15, France

Received 13 August 1993/Returned for modification 15 September 1993/Accepted 5 October 1993

The iota toxin which is produced by *Clostridium perfringens* type E, is a binary toxin consisting of two independent polypeptides: Ia, which is an ADP-ribosyltransferase, and Ib, which is involved in the binding and internalization of the toxin into the cell. Two degenerate oligonucleotide probes deduced from partial amino acid sequence of each component of *C. spiroforme* toxin, which is closely related to the iota toxin, were used to clone three overlapping DNA fragments containing the iota-toxin genes from *C. perfringens* type E plasmid DNA. Two genes, in the same orientation, coding for Ia (387 amino acids) and Ib (875 amino acids) and separated by 243 noncoding nucleotides were identified. A predicted signal peptide was found for each component, and the secreted Ib displays two domains, the propeptide (172 amino acids) and the mature protein (664 amino acids). The Ia gene has been expressed in *Escherichia coli* and *C. perfringens*, under the control of its own promoter. The recombinant polypeptide obtained was recognized by Ia antibodies and ADP-ribosylated actin. The expression of the Ib gene was obtained in *E. coli* harboring a recombinant plasmid encompassing the putative promoter upstream of the Ia gene and the Ia and Ib genes. Two residues which have been found to be involved in the NAD+ binding site of diphtheria and pseudomonas toxins are conserved in the predicted Ia sequence (Glu-14 and Trp-19). The predicted amino acid Ib sequence shows 33.9% identity with and 54.4% similarity to the protective antigen of the anthrax toxin complex. In particular, the central region of Ib, which contains a predicted transmembrane segment (Leu-292 to Ser-308), presents 45% identity with the corresponding protective antigen sequence which is involved in the translocation of the toxin across the cell membrane.

Clostridium perfringens is a ubiquitous bacteria which causes food-borne illness and gas gangrene in humans and digestive diseases in animals. This species is divided into five toxin types on the basis of the production of four major lethal toxins (alpha, beta, epsilon, and iota) (6). *C. perfringens* type E produces the iota toxin and has been implicated in calf and lamb enterotoxemias (6).

Iota toxin is a binary toxin which is composed of two independent polypeptide chains called iota a (Ia) (M_r, 47,500) and iota b (Ib) (M_r 71,500), which are not associated by either covalent or noncovalent bonds (26, 27). It has been shown that the Ia light chain causes ADP-ribosylation of globular skeletal muscle and nonmuscle actin at Arg-177 (23, 28). The heavy-chain Ib is required for penetration of Ia into the cytosol, and Ib must undergo limited proteolysis to be functionally active (3).

Iota toxin shares a comparable structure and mode of action with both *Clostridium spiroforme* toxin and C2 toxin of *Clostridium botulinum* types C and D (3). However, iota toxin is antigenically related to *C. spiroforme* toxin but not to *C. botulinum* C2 toxin (16, 19). An ADP-ribosyltransferase antigenically related to the enzymatic component of iota and *C. spiroforme* toxins was found in one *Clostridium difficile* strain (18). Iota and C2 toxins, like leucocidin (15), *Staphylococcus* gamma lysin (5), and anthrax (13) toxins, belong to the binary toxins formed of independent enzymatic and binding components. Therefore, this toxin family differs from the classical A-B toxins, such as the cholera and related toxins, which consist of a subunit or domain (A), with a specific function, and a binding domain or subunit(s) (B), and are assembled in a defined structure (25). In contrast, the protective antigen (PA) of the anthrax toxin binds to a cell surface receptor, is cleaved by a cell surface protease, and then is able to bind the enzymatic component (edema factor or lethal factor) (13).

Little is known about the genetics of the *Clostridium* binary toxins. In the present study, we report the characterization of the iota-toxin genes and the relationship with anthrax toxin genes. We also studied their expression in recombinant *Escherichia coli* and *C. perfringens*.

MATERIALS AND METHODS

C. spiroforme toxin protein sequencing. Light (Sa) and heavy chains (Sb) of *C. spiroforme* toxin were purified from *C. spiroforme* NCTC 11493 as previously described (19). Purified toxin chains were run on a 10% polyacrylamide gel containing 0.1% sodium dodecyl sulfate (SDS) and transferred to a polyvinylidene difluoride membrane (Immobilon; Millipore, Paris, France). After Coomassie blue staining and destaining, the protein bands were cut out and digested by trypsin. The peptides were separated by high-performance liquid chromatography, and one internal peptide of each toxin chain was microsequenced with a gas-phase protein sequencer (Applied Biosystems).

Bacterial DNA and plasmids. *C. spiroforme* NCTC 11493 and Cs246, and *C. perfringens* type E strain NCIB 10748 were grown in broth containing Trypticase (30 g/liter), yeast extract (20 g/liter), glucose (5 g/liter), and cysteine-HCl (0.5 g/liter) (pH 7.2) under anaerobic conditions. *Clostridium* genomic and plasmid DNAs were extracted and purified as previously described (17).

Plasmid pUC19 (Appligene, Strasbourg, France) and pACYC184 (New England Biolabs, Ozyme, Paris, France) were used for cloning in *E. coli* TG1 (*Δlac-pro*) *thi supE hsdD5* F' (*traD36 proAB*$^+$ *Δlac*ZM15), and pJIR418 (24) was used for expression in *C. perfringens* S13 (24).

* Corresponding author.

Probes and hybridization conditions. Oligonucleotides were synthesized by the phosphoramidite method with a Cyclone Milligen automated DNA synthesizer. Genescreen Plus filters (New England Nuclear Research Products, du Pont Nemours, Paris, France) were pretreated with 200 μg of heat-denaturated salmon sperm DNA per ml in 1 M NaCl–10% dextran sulfate–0.5% SDS–50 mM Tris-HCl (pH 7.5) at 40°C and then treated with a 5′-^{32}P-labeled oligonucleotide (10^6 cpm/ml) in the same mixture overnight at 40°C. Filters were washed in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at 40°C for 2 h and exposed overnight to Fujii RX films.

The recombinant DNAs used as probes were labeled by the random method with the Multiprime kit (Amersham, Paris, France). The hybridizations were carried out with the Rapid Hybridization Buffer (Amersham) at 65°C for 4 h. The filters were washed in 0.1× SSC–0.1% SDS at 65°C for 1 h.

PCR amplification. One hundred nanograms of DNA was amplified by the polymerase chain reaction (PCR) in a total volume of 100 μl of 10 mM Tris-HCl (pH 8.3)–50 mM KCl–4 mM $MgCl_2$–0.1% bovine serum albumin–100 μM deoxynucleoside triphosphate–10 mM β-mercaptoethanol–50 pmol of each primer–2.5 U of *Taq* polymerase (Amersham). Reaction mixtures were denatured at 95°C for 2 min and then subjected to 30 cycles of denaturation (20 s at 94°C), annealing (20 s at 5°C below the theoretical melting temperature of the primers), and extension (20 s at 72°C) in a Prem III thermal cycler (Flobio, Paris, France).

Other molecular biology techniques. Ligation and preparation of plasmid DNA from *E. coli* were conducted as described by Maniatis et al. (14). Bacteria were transformed by electroporation (17). *C. perfringens* electroporation was done according to the method of Scott and Rood (22). T4 polynucleotide kinase and calf intestinal phosphatase were obtained from Boehringer-Mannheim France, other enzymes were obtained from Pharmacia (Paris, France), and the Erase-a-base was obtained from Promega (Coger, Paris, France). DNA was sequenced by the dideoxy chain terminator procedure with a Sequenase kit (United States Biochemical Corp., Cleveland, Ohio).

Potential biohazards associated with the experiments described had been approved by the French National Control Committee.

Expression of recombinant proteins. *E. coli* TG1 transformed with recombinant plasmid pUC19 or pACYC184 was grown in LB medium (14)–50 μg of ampicillin per ml to an optical density at 600 nm of 1. One milliliter of the cell suspension was pelleted in a microcentrifuge, resuspended in 100 μl of 20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 7.5), and sonicated. One to 20 μl was mixed with 5 μl of Laemmli sample buffer, boiled for 3 min, and loaded onto a 0.1% SDS–12% polyacrylamide gel (12).

For immunobloting analysis, proteins were transferred electrophoretically to nitrocellulose (Hybond C; Amersham). The nitrocellulose was first incubated for 1 h in phosphate-buffered saline containing 5% milk and then incubated overnight at room temperature with 1:400 dilution of immunopurified Ia or Sb (18, 19) rabbit antibodies. Bound antibodies were detected with peroxidase-labeled protein A and the chemiluminescence kit provided by Amersham.

Gel assay for ADP-ribosylation. In vitro ADP-ribosylation assay were performed with G actin isolated from *Xenopus laevis* oocytes as described previously (18). The mixture for the polyacrylamide gel electrophoresis (PAGE) assay (total volume, 20 μl) contained 7 μg of oocyte actin, 100 mM

```
Sb   256 Y   L   S   S   Y   L   E   S   N   T   A
P131 5'-TAT TTI TCI TCI TAT TTI GAA TCI AAT ACI GCI
b       TAT GTA TCA AGT TAT TTA GAA TCA AAT ACC GCT

Sb       G   D   P   Y 270
P131    GGI GAT CCI TAT -3'
b       GGT GAC CCA TAT


Sa   127 D   K   P   M   Y   V   Y   Y   F   E   S
P222 5'-GAT AAA CCI ATG TAT GTI TAT TAT TTT GAA TCI
b       GAT AAA CCT ATA AAT GTT TAT TAT TTT GAG TCT

Sa       P   E 139
P222    CCI GAA-3'
b       CCA GAG
```

FIG. 1. Amino acid sequences of internal peptides of *C. spiroforme* toxin components Sb and Sa and oligonucleotide probes P131 and P222 complementary to these sequences, respectively, according to the *Clostridium* codon usage (see the text). The nucleotide sequences determined by DNA sequencing of *C. perfringens* type E (NCIB 10748) are shown in lines b. The amino acids are numbered according to the iota-toxin sequence.

HEPES, 5 mM ATP, 2.5 mM ADP-ribose, and 5 × 10^5 cpm of ^{32}P-NAD (specific activity, 30 Ci/mmol; Dupont, NEN Research Products, Boston, Mass.). Between 1 and 10 μl of a fraction to be tested for enzymatic activity was added. After incubation for 1 h at 37°C, sample buffer was added and the preparation was fractionated by SDS-PAGE and processed by autoradiography.

Purification of the recombinant Ia. *C. perfringens* S13 harboring pMRP76 was grown in TGY under anaerobic conditions to an optical density at 600 nm of 1. Culture supernatant (1 liter) was precipitated by ammonium sulfate (70% saturation). The precipitate was dissolved in 30 ml of 10 mM Tris (pH 7.5), dialyzed, and applied to a column (2 by 15 cm) containing DEAE-Sepharose CL6B (Pharmacia) equilibrated with 10 bed volumes of the same buffer. The gel was washed and eluted with 0.1 M NaCl in the same buffer. The eluate was purified on an immunoaffinity column with rabbit antibodies against *C. perfringens* toxin component Ia as previously described (19). The column was eluted with 4 M guanidine hydrochloride (Sigma) in 10 mM Tris (pH 7.5). The eluted fractions were dialyzed against the Tris buffer.

Nucleotide sequence accession number. The nucleotide sequences reported in this paper have been submitted to the EMBL Data Library with accession number X73562.

RESULTS

Cloning of the iota-toxin genes. Several different peptides of *C. spiroforme* toxin chains Sa and Sb were microsequenced. Two oligonucleotides (P131 and P222), deduced from internal peptide sequences of Sb and Sa, respectively, were synthesized according to the *Clostridium* codon usage (31) and with inosine at the most degenerated positions (Fig. 1). These probes hybridized with total and plasmid DNA of *C. spiroforme* NCTC 11493 and Cs246 and *C. perfringens* NCIB 10748 (data not shown), indicating that iota-toxin genes are localized on plasmid DNA of these *Clostridium* strains. Since DNA was easier to prepare from *C. perfringens* than from *C. spiroforme*, we chose to clone iota-toxin genes from *C. perfringens* NCIB 10748.

The 2-kb *Hin*dIII DNA fragment of *C. perfringens* NCIB 10748 plasmid DNA recognized by P131 was cloned into

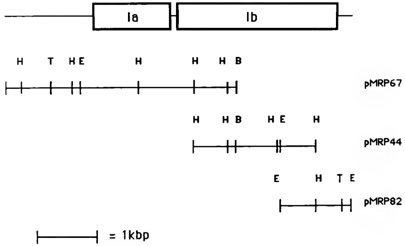

FIG. 2. Cloning strategy of the iota-toxin genes encoding the enzymatic Ia components and binding Ib components. B, *Bcl*I; E, *Eco*RI; H, *Hin*dIII; T, *Taq*I.

pUC19 (pMRP44). The synthetic oligonucleotide P190 complementary to the 5′ extremity of the 2-kb *Hin*dIII DNA fragment (nucleotides 3288 to 3308) was used to clone the 3.8-kb *Bcl*I overlapping DNA fragment (pMRP67) (Fig. 2). A third clone (pMRP82), which contains the 1.2-kb *Eco*RI DNA fragment overlapping with the 3′ extremity of the insert of pMR44, was obtained by using the primer P238 complementary to nucleotides 5141 to 5161.

The primer P222, deduced from the sequence of an internal peptide of Sa, hybridized with *C. spiroforme* and *C. perfringens* type E DNAs and also with pMRP67, suggesting that the Ia gene is in the proximity of the Ib gene.

The overlapping DNA fragments cloned in pMRP44, pMRP67, and pMRP82 were sequenced (Fig. 3). The DNA fragment cloned in pMRP44 was found to contain two internal *Hin*dIII sites, indicating a partial digestion of the DNA used for cloning or a rearrangement of the cloned *Hin*dIII DNA fragments. One of the internal *Hin*dIII sites was localized on the overlapping region with pMRP67 (Fig. 2 and 3). Two primers were synthesized on each part of the second *Hin*dIII site and in the opposite direction from P231 (positions 4638 to 4618) and P212 (positions 4225 to 4241) (Fig. 3). The PCR amplification products with these primers, *C. perfringens* DNA, and pMRP44 were the same size (data not shown). These results showed that the two internal *Hin*dIII sites in pMRP44 are the result of a partial DNA digestion.

Features of iota-toxin genes. The 5,747 nucleotide sequence showed two open reading frames. The open reading frame near the 5′ extremity and recognized by P222, which was deduced from the Sa internal peptide sequence, was assigned to the Ia gene, and the downstream open reading frame, recognized by P131, was deduced from the internal Sb sequence and assigned to the Ib gene.

The Ia gene starts at the initiation codon ATG at position 1465 and ends at the stop codon at position 2628. A consensus ribosome binding site, GGAGGG, is localized 6 nucleotides upstream of the initiation codon. DNA stretches TATAAT and TTGTCAT, homologous to the −10 and −35 *Clostridium* consensus promoter regions (31), are identified between positions 1304 and 1333 (Fig. 3). Three direct repeats and one inverted repeat are localized in the promoter region (Fig. 3).

The Ib gene, from the initiation codon ATG (position 2872) to the stop codon at position 5499, is preceded by a consensus ribosome binding site (GGAGG) between positions 2860 and 2864. The Ib gene is located on the same frame as the Ia gene. On the 243 noncoding nucleotides between the two genes, no promoter consensus sequence has been found. No long inverted repeats have been identified downstream of the Ia and Ib genes.

Analysis of the deduced amino acid sequences. (i) Ia sequence. The deduced polypeptide is composed of 387 amino acids and has a predicted molecular mass of 45,121 Da. The 41 N-terminal amino acids contained residues characteristic of a signal peptide (20): the presence of charged N-terminal residues at positions 2, 3, and 6 followed by a long hydrophobic core, of a turn residue (Pro-23), and of a proteolytic cleavage site (Arg-41 to Ala-42). Furthermore, the amino acid sequence from Ala-42 to Glu-56 is in agreement with the sequence determined by protein sequencing of the N-terminal part of Sa (X-X-Ile-Glu-Pro-Pro-Pro-Asp-Phe-Leu-Lys-Lys-Lys-Glu). The discordant residues could be due to differences between Sa and Ia or to inaccuracies in protein sequencing (the presence of contaminating proteins or breakdown products which interfere with sequence analysis). However, the cleavage sites of the clostridial signal peptidases have not been fully investigated. It is not excluded that the cleavage by signal peptidase occurs at Ala-42, or Ala-34 with a subsequent cleavage at Arg-41 to Ala-42 by clostridial peptidase, since Ala-X represents a common site of bacterial signal peptidases (29).

The mature Ia protein is formed of 346 amino acids. The predicted molecular mass (40,392 Da) is comparable to that found for purified native Sa (44,000) and slightly lower than that of native Ia (47,500). The predicted pI of 5.01 agrees with the experimentally determined pI of Ia (5.2) and is lower than that of Sa (pI 6.2) (19, 26). The differences in the predicted and experimentally determined molecular masses of Ia could be due to a particular conformational structure.

The Ia sequence did not show any significant similarity with other known protein sequences by FASTA and Blast programs. Domenighini et al. have identified conserved amino acids involved in the NAD+ binding site of ADP-ribosylating toxins (4). Recently, it has been shown that the peptide Phe-9 to Gly-19 of the ADP-ribosylating enzyme C3

```
1    GATCATTTACTCCATCACCTAAAGACTCTTGTAATTTCACTAGATTATTAAAGTTATTTGAACA
65   AATATTAAAATCAATAATTTTACTAGCCCAAATTCTATATTTTTGTACAGTAGCTGTAATTATT
129  AAGTGTAAATTAAGCTTTTGAGTCCAAGGATATGAGTTAGAAGAACACCAATCATATAAATTAC
193  CACATGGATTTGAGTTTAATGCATAAGTTAAATTTAATCTTTCTGTAGAAGTAGCTAAATCAAG
257  AATTTCTTTTTATATCTGAAAATGGAACTATTAAAACACATCTTTTACCAACTTCAATTAATGA
321  ACCTTGATTAATATTTTATAAGCTATCTGCAGATGTTTTACTAAGTCATATTCCAGTCAATTTA
385  TAAATTGCTCTATCATCAGATATACTACCACGAGAAATTCTAATAGCTTGATACTTTCTATAAG
449  TTGCATAAACCTTATAATATACATACTCATTTGGAACAGTAGTTGTAGATACAGATCTTTCTAT
513  TGTATTTTGTTACCCTATAGTTATTCAAAATCCATATTCTACAGATGCTTGAATAAATTGATAA
577  GTAAAGCCAACTGAAAAGTTTACATTTATAGATACTTCTTTAGATTTAGTTAGTGATTGAATAA
641  AAGTATATGTTTAATTAAGATTAAGAATTTGACTTGAAACTACAGAAGGTTTACTTAATATCCA
705  ACCATCTCCTTTATCTATTACATATAATCAATCACTTAAATTCGAGTTAGAGTTTTCAAGATTA
769  ATTGGTATTTTTAAATCCTAAGAAACAAGCAATACTTCTTTAGCATCTTCAAACACCATTGGAT
833  TTAAATTAGTATTAAATATATTTGTGAAATTTTTCATATAAATAGAAATATTTAATTTTTGATA
897  ATTTAAAGAATAGAAAAAGAGAACTTAGTACTATTAATAATACTGATGAAATTTTAAAAACTAG
961  AATACTTTTAAATTTATATATTAACTATAATCTTTTGATATATTATCTGGATAATAGCATATTT
1025 AAAGCTTATCTAATAACTTATTATGAAAGAAAAGACTTTAATAAGAATATAAATACTTTATTTG
1089 TTTACAAATATCAATTATAAAATCTTTACTTCTATACATATTAATAATTTATATTCATTTAGAT
1153 AGAAAATCAGATTATTGTTTAAAGTTTTGAATTTGTTTTAATAAATTTTAATTATTATAAATAT
        EcoRI
1217 AGAGAATTCAGAAAATACAATCTAATTTTACATTAAATTACTTTTATTGATAAAAATATATTTT
                                   -35                    -10
1281 TTGTATTTATATAGAGATTTTCTTTGTCATATACTGTATAATATTTATATAATATTAATACATA
                                             ------->  ------->     ---
                                                      <---------     --

1345 ATATTATAACTAAATTCAAAATAGAAAGGAGGTTATCTTATGTTAAAAAGAAAAATTTAAAATA
     ---->
     ------->
                                                        S.D.  Ia--> M   K
1409 TATTTTCATTATTTGAAGATTGAATCTTAAATAAACTTGAATTAGGAGGGAGTATT ATG AAA

   3 K   V   N   K   S   I   S   V   F   L   I   L   Y   L   I   L
1471 AAA GTT AAT AAA TCT ATA TCT GTA TTT CTA ATA TTA TAT TTA ATT TTA

  19 T   S   S   F   P   S   Y   T   Y   A   Q   D   L   Q   I   A
1519 ACT AGT TCA TTT CCT AGT TAT ACT TAT GCA CAA GAT TTA CAA ATA GCA

  35 S   N   Y   I   T   D   R   A   F   I   E   R   P   E   D   F
1567 AGC AAT TAT ATT ACA GAT AGA GCT TTT ATT GAA AGA CCA GAA GAT TTT

  51 L   K   D   K   E   N   A   I   Q   W   E   K   K   E   A   E
1615 CTT AAA GAT AAA GAA AAT GCT ATT CAA TGG GAA AAA AAG GAG GCT GAA

  67 R   V   E   K   N   L   D   T   L   E   K   E   A   L   E   L
1663 AGA GTA GAA AAA AAC CTT GAT ACA CTT GAA AAA GAA GCA TTA GAA TTA

  83 Y   K   K   D   S   E   Q   I   S   N   Y   S   Q   T   R   Q
1711 TAT AAA AAA GAT TCT GAA CAA ATA AGT AAC TAC TCT CAG ACA AGA CAG

  99 Y   F   Y   D   Y   Q   I   E   S   N   P   R   E   K   E   Y
1759 TAT TTT TAC GAC TAT CAA ATA GAA TCA AAT CCT AGA GAA AAA GAA TAC

 115 K   N   L   R   N   A   I   S   K   N   K   I   D   K   P   I
1807 AAA AAT CTT AGA AAT GCC ATA TCA AAA AAT AAG ATA GAT AAA CCT ATA

 131 N   V   Y   Y   F   E   S   P   E   K   F   A   F   N   K   E
1855 AAT GTT TAT TAT TTT GAG TCT CCA GAG AAA TTT GCG TTT AAT AAA GAA

 147 I   R   T   E   N   Q   N   E   I   S   L   E   K   F   N   E
1903 ATA AGA ACA GAA AAT CAA AAT GAA ATT TCT TTA GAG AAA TTT AAT GAG

 163 L   K   E   T   I   Q   D   K   L   F   K   Q   D   G   F   K
1951 TTG AAA GAA ACT ATT CAA GAT AAA TTG TTT AAA CAA GAT GGA TTT AAG

 179 D   V   S   L   Y   E   P   G   N   G   D   E   K   P   T   P
1999 GAT GTT TCT TTA TAT GAA CCA GGT AAT GGC GAT GAA AAG CCT ACA CCA

 195 L   L   I   H   L   K   L   P   K   N   T   G   M   L   P   Y
2047 CTA CTT ATA CAT TTG AAA TTA CCA AAA AAT ACT GGT ATG TTA CCA TAT

 211 I   N   S   N   D   V   K   T   L   I   E   Q   D   Y   S   I
2095 ATA AAT TCT AAT GAT GTA AAA ACA TTA ATA GAA CAA GAC TAT AGC ATA

 227 K   I   D   K   I   V   R   I   V   I   E   G   K   Q   Y   I
2143 AAG ATA GAC AAA ATT GTT CGT ATA GTA ATA GAA GGA AAG CAA TAT ATA

 243 K   A   E   A   S   I   V   N   S   L   D   F   K   D   D   V
2191 AAA GCT GAA GCT TCT ATT GTA AAC AGT CTT GAT TTT AAA GAT GAT GTA

 259 S   K   G   D   L   W   G   K   E   N   Y   S   D   W   S   N
2239 AGT AAA GGT GAT TTA TGG GGA AAA GAA AAT TAT AGT GAT TGG AGT AAT

 275 K   L   T   P   N   E   L   A   D   V   N   D   Y   M   R   G
2287 AAA TTA ACT CCT AAT GAA CTT GCT GAT GTA AAT GAC TAT ATG CGT GGA

 291 G   Y   T   A   I   N   N   Y   L   I   S   N   G   P   L   N
2335 GGA TAT ACC GCA ATT AAT AAC TAT TTA ATA TCA AAT GGT CCT TTA AAT

 307 N   P   N   P   E   L   D   S   K   V   N   N   I   E   N   A
2383 AAT CCT AAT CCA GAA CTA GAC TCT AAA GTA AAT AAC ATT GAA AAC GCA

 323 L   K   L   T   P   I   P   S   N   L   I   V   Y   R   R   S
2431 TTA AAG CTC ACA CCT ATT CCA TCT AAC TTA ATT GTA TAT AGA AGG TCT

 339 G   P   Q   E   F   G   L   T   L   T   S   P   E   Y   D   F
2479 GGT CCA CAA GAA TTT GGA TTA ACT CTC ACA TCT CCT GAA TAT GAT TTT

 355 N   K   I   E   N   I   D   A   F   K   E   K   W   E   G   K
2527 AAT AAA ATA GAA AAT ATA GAT GCT TTT AAA GAA AAA TGG GAA GGA AAA

 371 V   I   T   Y   P   N   F   I   S   T   S   I   R   K   C   K
2575 GTA ATA ACA TAC CCA AAC TTT ATT AGT ACT AGT ATT CGG AAG TGT AAA

 387 I   *
2623 ATA TGA GTGCATTTGCTAAAAGAAAAATAATACTACGTATAAACATACCAAAAGATTCTCCA
2685 GGAGCTTATTTATCAGCCATTCCAGGTTATGCAGGAGAATATGAAGTACTTTTAAATCATGGAA
2749 GTAAATTTAAAATCAATAAAGTTGATTCTTATAAAGATGGAACTGTAACAAAACTAATTTTGGA
   1                                                      S.D.   Ib--> M
2813 TGCAACATTGATAAATTAATACTTTTAATATAAATAATCAATTTAAAGGAGGAAAAAT A ATG

   2 N   I   Q   I   K   N   V   F   S   F   L   T   L   T   A   M
2875 AAT ATA CAA ATT AAA AAT GTA TTT AGT TTT TTA ACA CTT ACA GCT ATG

  18 I   S   Q   T   L   S   Y   N   V   Y   A   Q   T   T   T   Q
2923 ATA AGT CAA ACG TTA TCA TAT AAT GTA TAT GCA CAA ACT ACT ACA CAA

  34 N   D   T   N   Q   K   E   E   I   T   N   E   N   T   L   S
2971 AAT GAT ACC AAT CAG AAA GAA GAA ATA ACA AAT GAA AAT ACA CTA TCA

  50 S   N   G   L   M   G   Y   Y   F   A   D   E   H   F   K   D
3019 AGT AAC GGA TTA ATG GGA TAT TAT TTT GCA GAT GAA CAT TTT AAG GAT

  66 L   E   L   M   A   P   I   K   N   G   D   L   K   F   E   E
3067 TTA GAA TTA ATG GCG CCA ATC AAA AAT GGT GAT TTG AAA TTT GAA GAA

  82 K   K   V   D   K   L   L   T   E   D   N   S   S   I   K   S
3115 AAG AAA GTA GAC AAG CTT TTA ACT GAA GAT AAT TCA AGT ATA AAA TCT

  98 I   R   W   T   G   R   I   I   P   S   E   D   G   E   Y   I
3163 ATC CGA TGG ACA GGA AGG ATA ATT CCT TCT GAA GAT GGT GAA TAT ATA

 114 L   S   T   D   R   N   D   V   L   M   Q   I   N   A   K   G
3211 TTG TCA ACT GAT AGA AAT GAT GTA TTA ATG CAA ATA AAT GCT AAA GGG

 130 D   I   A   K   T   L   K   V   N   M   K   K   G   Q   A   Y
3259 GAT ATT GCA AAA ACA CTT AAG GTT AAT ATG AAA AAA GGT CAG GCA TAC

 146 N   I   R   I   E   I   Q   D   K   N   L   G   S   I   D   N
3307 AAT ATT AGA ATA GAA ATA CAA GAC AAA AAT TTA GGT TCA ATT GAT AAT

 162 L   S   V   P   K   L   Y   W   E   L   N   G   N   K   T   V
3355 TTA TCT GTT CCT AAA CTT TAT TGG GAA TTA AAT GGA AAT AAA ACA GTA

 178 I   P   E   E   N   L   F   F   R   D   Y   S   K   I   D   E
3403 ATA CCT GAA GAA AAC TTA TTT TTC CGA GAT TAC TCT AAA ATA GAT GAA

 194 N   D   P   F   I   P   N   N   N   F   F   D   V   R   F   F
3451 AAT GAT CCG TTT ATA CCT AAT AAT AAT TTT TTT GAT GTA AGA TTT TTT

 210 S   A   A   W   E   D   E   D   L   D   T   D   N   D   N   I
3499 AGC GCA GCC TGG GAA GAT GAA GAT TTA GAT ACT GAT AAT GAT AAT ATT

 226 P   D   A   Y   E   K   N   G   Y   T   I   K   D   S   I   A
3547 CCA GAT GCT TAT GAA AAA AAT GGC TAT ACT ATC AAA GAT TCA ATT GCA

 242 V   K   W   N   D   S   F   A   E   Q   G   Y   K   K   Y   V
3595 GTA AAA TGG AAT GAT AGT TTT GCA GAA CAA GGA TAT AAA AAA TAT GTA

 258 S   S   Y   L   E   S   N   T   A   G   D   P   Y   T   D   Y
3643 TCA AGT TAT TTA GAA TCA AAT ACC GCT GGT GAC CCA TAT ACA GAT TAT

 274 Q   K   A   S   G   S   I   D   K   A   I   K   L   E   A   R
3691 CAA AAA GCT TCA GGT TCT ATT GAT AAA GCT ATA AAA TTA GAA GCA AGA

 290 D   P   L   V   A   A   Y   P   V   V   G   V   G   M   E   N
3739 GAT CCT TTA GTT GCA GCA TAT CCA GTT GTT GGA GTA GGT ATG GAA AAT

 306 L   I   I   S   T   N   E   H   A   S   S   D   Q   G   K   T
3787 TTA ATT ATA TCT ACT AAT GAA CAT GCT TCA AGT GAT CAA GGA AAA ACA

 322 V   S   R   A   T   T   N   S   K   T   D   A   N   T   V   G
3835 GTT TCT AGG GCT ACT ACA AAT AGT AAA ACT GAT GCA AAT ACA GTT GGA

 338 V   S   I   S   A   G   Y   Q   N   G   F   T   G   N   I   T
3883 GTA TCT ATT AGT GCT GGA TAT CAA AAT GGA TTT ACT GGT AAT ATA ACT

 354 T   S   Y   S   H   T   T   D   N   S   T   A   V   Q   D   S
3931 ACA AGC TAT TCT CAC ACA ACA GAT AAT TCA ACT GCT GTG CAA GAT AGT

 370 N   G   E   S   W   N   T   G   L   S   I   N   K   G   E   S
3979 AAT GGA GAA TCA TGG AAT ACT GGA TTA AGT ATA AAT AAA GGA GAA TCA

 386 A   Y   I   N   A   N   V   R   Y   Y   N   T   G   T   A   P
4027 GCA TAT ATA AAT GCC AAT GTA AGA TAT TAT AAT ACT GGT ACT GCA CCT

 402 M   Y   K   V   T   P   T   T   N   L   V   L   D   G   E   T
```

FIG. 3. Nucleotide sequence and amino acid translation of the iota-toxin genes. The putative Shine Dalgarno (S.D.) and −35 and −10 promoter sequences are underlined. The predicted signal peptides are italicized. Stop codons are indicated by an asterisk, and the directed and inverted repeats on the promoter region are indicated by dashed arrows beneath the sequence.

```
4075 ATG TAC AAA GTA ACT CCA ACA ACC AAT TTA GTA TTA GAT GGA GAG ACA

 418 L   A   T   I   K   A   Q   D   N   Q   I   G   N   N   L   S
4123 TTA GCA ACT ATT AAA GCA CAG GAT AAT CAA ATT GGT AAT AAC TTA TCT

 434 P   N   E   T   Y   P   K   K   G   L   S   P   L   A   L   N
4171 CCA AAT GAA ACA TAT CCT AAA AAA GGA CTT TCT CCT TTA GCT CTT AAC

 450 T   M   D   Q   F   N   A   R   L   I   P   I   N   Y   D   Q
4219 ACA ATG GAT CAA TTT AAT GCT AGG TTA ATT CCA ATA AAT TAC GAT CAA

 466 L   K   K   L   D   S   G   K   Q   I   K   L   E   T   T   Q
4267 CTT AAA AAA TTA GAT TCT GGA AAA CAA ATT AAA TTA GAA ACA ACA CAA

 482 V   S   G   N   Y   G   T   K   N   S   Q   G   Q   I   I   T
4315 GTA AGT GGA AAT TAT GGA ACT AAA AAT AGT CAA GGA CAA ATA ATT ACA

 498 E   G   N   S   W   S   N   Y   I   S   Q   I   D   S   V   S
4363 GAA GGA AAT AGT TGG TCT AAC TAT ATA AGT CAA ATT GAT AGC GTT TCT

 514 A   S   I   I   L   D   T   G   S   Q   T   F   E   R   R   V
4411 GCA TCT ATT ATA CTG GAT ACA GGA AGT CAA ACA TTT GAA AGG CGT GTG

 530 A   A   K   E   Q   G   N   P   E   D   K   T   P   E   I   T
4459 GCT GCT AAA GAG CAA GGA AAT CCA GAA GAT AAA ACT CCT GAG ATT ACA

 546 I   G   E   A   I   K   K   A   F   S   A   T   K   N   G   E
4507 ATT GGA GAA GCA ATT AAA AAA GCT TTT AGT GCT ACT AAA AAT GGT GAA

 562 L   L   Y   F   N   G   I   P   I   D   E   S   C   V   E   L
4555 TTA TTA TAT TTT AAT GGA ATT CCA ATT GAT GAG AGC TGT GTT GAA CTT

 578 I   F   D   D   N   T   S   E   I   I   K   E   Q   L   K   Y
4603 ATA TTT GAT GAT AAT ACA TCT GAA ATA ATT AAA GAA CAA TTA AAA TAT

 594 L   D   D   K   K   I   Y   N   V   K   L   E   R   G   M   N
4651 TTA GAT GAT AAA AAG ATA TAT AAT GTT AAA CTT GAA AGA GGA ATG AAT

 610 I   L   I   K   V   P   S   Y   F   T   N   F   D   E   Y   N
4699 ATA CTT ATA AAG GTG CCT TCA TAT TTT ACT AAT TTT GAT GAA TAT AAT

 626 N   F   P   A   S   W   S   N   I   D   T   K   N   Q   D   G
4747 AAT TTT CCT GCT TCA TGG AGT AAT ATT GAT ACT AAA AAC CAA GAT GGT

 642 L   Q   S   V   A   N   K   L   S   G   E   T   K   I   I   I
4795 TTA CAA AGT GTA GCA AAT AAG TTA AGC GGA GAG ACA AAG ATT ATA ATA

 658 P   M   S   K   L   K   P   Y   K   R   Y   V   F   S   G   Y
4843 CCT ATG TCT AAA TTA AAA CCC TAT AAA CGC TAT GTT TTT AGT GGA TAT

 674 S   K   D   P   S   T   S   N   S   I   T   V   N   I   K   S
4891 TCA AAG GAT CCT TCA ACT TCT AAT TCA ATA ACA GTA AAT ATA AAA TCA

 690 K   E   Q   K   T   D   Y   L   V   P   E   K   D   Y   T   K
4939 AAA GAA CAG AAA ACA GAT TAT TTA GTA CCA GAG AAA GAT TAT ACA AAA

 706 F   S   Y   E   F   E   T   T   G   K   D   S   S   D   I   E
4987 TTT AGT TAT GAA TTT GAA ACA ACC GGA AAA GAT TCT TCT GAT ATA GAA

 722 I   T   L   T   S   S   G   V   I   F   L   D   N   L   S   I
5035 ATA ACA TTA ACA AGT AGT GGT GTA ATA TTT TTA GAT AAT TTA TCT ATT

 738 T   E   L   N   S   T   P   E   I   L   K   E   P   E   I   K
5083 ACA GAA TTA AAT AGT ACT CCT GAA ATA TTA AAA GAA CCA GAA ATT AAA

 754 V   P   S   D   Q   E   I   L   D   A   H   N   K   Y   Y   A
5131 GTT CCA AGT GAC CAA GAA ATA CTA GAT GCA CAT AAC AAA TAT TAT GCA

 770 D   I   K   L   D   T   N   T   G   N   T   Y   I   D   G   I
5179 GAT ATA AAG CTT GAC ACA AAT ACA GGA AAC ACT TAT ATA GAT GGT ATA

 786 Y   F   E   P   T   Q   T   N   K   E   A   L   D   Y   I   Q
5227 TAT TTT GAA CCA ACT CAA ACT AAT AAA GAA GCT CTT GAT TAT ATT CAA

 802 K   Y   R   V   E   A   T   L   Q   Y   S   G   F   K   D   I
5275 AAA TAT AGA GTT GAA GCA ACT TTG CAA TAT TCA GGA TTT AAA GAT ATT

 818 G   T   K   D   K   E   I   R   N   Y   L   G   D   Q   N   Q
5323 GGA ACT AAG GAT AAA GAA ATA CGT AAT TAT TTA GGA GAT CAA AAC CAA

 834 P   K   T   N   Y   I   N   F   R   S   Y   F   T   S   G   E
5371 CCT AAA ACT AAT TAT ATT AAT TTT AGA AGT TAT TTT ACT AGT GGA GAA

 850 N   V   M   T   Y   K   K   L   R   I   Y   A   V   T   P   D
5419 AAT GTT ATG ACA TAT AAA AAA TTA AGA ATA TAT GCA GTT ACA CCT GAT

 866 N   R   E   L   L   V   L   S   V   N   *
5467 AAT AGA GAG TTA TTA GTG CTT AGT GTT AAT TAA TACTAAATAAAAATTATATT

5520 TAAAATTATTAATAAATATCAAAATTAATAAAGCTAGTTATTCATACTGGCTACCTATATAAAT
5584 ATTTCTCGATTACTTACAGAGGATGTCAATAATTTTGTGTAAACTATCTAAAGAAATATATTGA
5648 AATTTGTTGAGGTTTAGATAATATTTTATTTAAACCTTAATTTTCTTTTTATGTAATTATAGTG
5712 AATATAAATTGTACAAAAACTAGAGATATTGAATTC
```

FIG. 3—*Continued.*

from *C. botulinum* C-003-9 constitutes the adenine ring-binding domain of the NAD+ binding site (1). In this peptide, Glu-13 and Trp-18 align with Glu-148 and Trp-153 of diphtheria toxin (DT), Glu-553 and Trp-558 of pseudomonas exotoxin A (ET), and Trp-26 of pertussis toxin (PT), which have been assigned to cross-link the nicotinamide ring of NAD+ (4). In the C3 enzyme of *C. botulinum* C-468 (17) and the epidermal cell differentiation inhibitor (EDIN) of *Staphylococcus aureus* (8), which is a related enzyme, Glu-13 is replaced by the equivalent residue Asp-13 (Fig. 4). By comparing the Ia sequence with those of the other ADP-ribosylating toxins, Glu-14 and Trp-19 of Ia can be aligned with the above conserved residues (Fig. 4). As for C3, EDIN, and PT, these conserved residues in Ia are localized near the N terminus.

```
Iota           10-L K D K E N A I Q W E K-21
DT            144-S S S V E Y I N N W E Q-155
ET            549-G G R L E T I L G W G N-560
PT             17-D V F Q N G F T A W G N-28
C3 (C-003-9)    9-F T N V E E A K K W G N-20
C3 (C-468)      9-F T N I D Q A K A W G N-20
EDIN            9-F T D L D E A T K W G N-20

consensus            E/D X X X X W
```

FIG. 4. Alignment of the Ia amino acid sequence with those involved in the NAD+ binding site of DT, ET, PT, C3 enzyme from *C. botulinum* C-003-9, from *C. botulinum* C-468, and from *Staphylococcus* EDIN. The conserved amino acids which have been assigned to cross-link the nicotinamide ring of NAD+ in DT, ET, PT, and the corresponding residues in the other sequences are shown in boldface.

(ii) Ib sequence. The Ib 875 deduced amino acids have a predicted molecular mass of 98,467 Da. The N-terminal 39 amino acids form an hydrophobic domain with an N-terminal-charged residue (Lys-6) and probably correspond to a peptide signal. A potential cleavage site for signal peptidase (Lys-39 to Glu-40) is localized at the extremity of this putative signal peptide. The predicted precursor protein consists of 836 residues (94,013 Da) and is in agreement with that found for the Sb precursor (92,000 Da) (19). The N-terminal sequence of the naturally activated form of Sb determined by protein sequencing (Gly-Trp-Gly-Asp-Glu-Asp-Leu-Asp) matches amino acids Ala-212 to Asp-219. The two unrelated residues could correspond to differences in Ib and Sb sequences or inaccuracies in protein sequencing. The deduced polypeptide from Ala-212 to Asn-875 (664 amino acids) is predicted to correspond to the functional proteolytically activated form of Ib or mature Ib. The predicted molecular mass (80,890 Da) and pI (4.67) are in agreement with those determined experimentally for Sb (76,000 Da and pI 4.7) (19) and differ slightly from those found for Ib (67,000 Da and pI 4.2) (26).

The propeptide (Glu-40 to Ala-211), which is excised from the Ib precursor, consists of 172 amino acids, with a predicted molecular mass of 19,883 Da and a pI of 4.62.

Similarity between Ib and PA. By comparison with protein sequences available in the data bank, Ib displays significant homology with only the PA of the anthrax toxin complex (13). Overall 33.9% identity and 54.4% similarity have been found between both protein sequences. The central regions of these proteins show the maximum homology, 45% identity and 70% similarity between amino acids 268 to 531 of Ib and amino acids 260 to 536 of PA.

Ib and PA present similar structures. Both contain a signal peptide (39 residues for Ib and 29 residues for PA), a propeptide (19,883 Da for Ib and 19,243 Da for PA), and a mature protein (80,000 Da for Ib and 63,000 Da for PA) (13).

```
        .   ┌─► IB propeptide      .          .         .
 40 .....EEITNENTLSSNGLMGYYFADEHFKDLELMAPIKNGDLKFEEKKV
         : :  ||..  ||.||:||||.|  :|..   ::..  ..|||.:....:
 30 EVKQENRLLNESESSSQGLLGYYFSDLNFQAPMVVTSSTTGDLSIPSSEL
    └─► PA20        .         .         .         .
 85 DKLLTEDNSSIKSIRWTGRIIPSEDGEYILST.DRNDVLMQINAKGDIAK
    :. :..:|   :.|   |.| |  ....:||.:.|  . |.| |  ::..:  |.|
 80 EN.IPSENQYFQSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINK

              .         .         .         .         .
134 T...LKVNMKKGQAYNIRIEIQDKNLGSIDNLSVPKLYWELNGNKTVIPE
    .    |:.:.||. |.|:|:.| .| .. ..|... .: : ...|.||..
129 ASNSNKIRLEKGRLYQIKIQYQREN.PTEKGLDFKLYWTDSQNKKEVISS
              .         .         .  ┌─► mature IB
181 ENLFFRDYSKIDENDPFIPNNNFFDVRFFSAAWEDEDLDTDNDNIPDAYE
    :|| :.:... ..|..            . |  ||:   ..   | |||.|||..|
178 DNLQLPELKQKSSNSR.........KKRSTSAG..PTVPDRDNDGIPDSLE
              .         .          └─► PA 63    .         .
231 KNGYTIKDSIAVKWNDSF.....AEQGYKKYVSSYLESNTAGDPYTDYQK
     :|||:.  .   .: ..:        ...|..|| ||   ...||:|||.|::|
218 VEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYSDFEK
              .         .         .ATP/GTP binding site
276 ASGSIDKAIKLEARDPLVAAYPVVGVGMENLIISTNEHASSD....QGKT
    ..|.|||.:.  |||.||||||||:|  |:|||:|:|.||..|.:     :.:|
268 VTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRT

              .         .         .         .         .
322 VSRATTNSKTDANTV..GVSISAGYQNGFTGNITTSYSHTTDNSTAVQDS
    :|:.|..|:|....|   ...: |::   ::.|.:..::|:......|::.|
318 ISKNTSTSRTHTSEVHGNAEVHASF.FDIGGSVSAGFSNSNSSTVAIDHS

              .         .         .         .         .
370 ....NGESWNTGLSINKGESAYINANVRYYNTGTAPMYKVTPTTNLVLD.
        .: .|...:::|.::.| :|||:|| ||||||:|.| |||.|||:
367 LSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYNVLPTTSLVLGK

              .         .         .         .         .
415 GETLATIKAQDNQIGNNLSPNETYPKKGLSPLALNTMDQFNARLIPINYD
    .:|||||||.:||::. |.||: ||.|.|.|:|||. |:|..   |.:||:
417 NQTLATIKAKENQLSQILAPNNYYPSKNLAPIALNAQDDFSSTPITMNYN

              .         .         .         .         .
465 QLKKLDSGKQIKLETTQVSGNYGTKNSQGQ..IITEGNSWSNYISQIDSV
    |: .|:..||::|:|.|| ||.:| | :..     :..|..||: :.||:..
467 QFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDTGSNWSEVLPQIQET

              .         .         .         .         .
513 SASIILDTGSQTF.ERRVAAKEQGNPEDKT.PEITIGEAIKKAFSATKNG
    .|.||::.  . .: |||:|| :.::| :.| |::|: ||:| ||: .. .
517 TARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPN

              .         .         .         .         .
561 ELLYFNGIPIDESCVELIFDDNTSEIIKEQLKYLDDKKIYNV....KLER
    : | :.| .|.|  .:: ||:.||: ||:||  |:...||.|     ||:
567 GNLQYQGKDITE..FDFNFDQQTSQNIKNQLAELNATNIYTVLDKIKLNA

              .         .         .         .         .
607 GMNILIKVPSYFTNFDEYNNFPASWSNIDTKNQDGLQSVANKLSGETKII
     |||||:                ..|. .:.||.. .:::: . |::   :...
615 KMNILIRD..........KRFHYDRNNIAVGADESVVKEAHREVINSSTE

              .         .         .         .         .
657 IPMSKLKPYKRYVFSGYSKDPSTSNSIT.........VNIKSKEQKTDYL
      : .:..  | ::|||  : ...:::.            :||.|  |... :
655 GLLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNISSLRQDGKTF

              .         .         .         .         .
698 VPEKDYTKFSYEFETTGKDSSDIEITLTSSGVIFLDNLSITELNSTPEIL
    :   |:.|:             : .:.| |.. :  |:          .:
705 I...DFKKY............NDKLPLYISNPNYKVNVY........AVT

              .         .         .         .         .
748 KEPEIKVPSDQEILDAHNKYYADIKLDTNTGNTYIDGIYFEPTQTNKEAL
    || .|  ||:.                     |:| .:||        |. |
732 KENTIINPSEN...................GDTSTNGI........KKIL
              .         .         .         .         .
798 DYIQK.YRVE
     : .| | ::
755 IFSKKGYEIG
```

FIG. 5. Alignment of Ib and PA (top and bottom lines, respectively, of each pair of lines) amino acid sequences. The numbering begins at the start of the signal peptide of Ib and PA. Arrows indicate the start of the propeptide of Ib and PA (PA20) and the start of mature Ib and PA (PA63). The furin cleavage site of PA is underlined. The predicted transmembrane segments are in boldface, and the predicted Ib ATP/GTP binding site is indicated by a dashed line. Sequences were aligned by the GAP program of the Genetics Computer Group, University of Wisconsin. Symbols vertically connecting homologous residues in the two sequences are "|" for identical residues, ":" for highly similar residues, and "." for less similar residues.

It has been shown that the cleavage between the propeptide and mature PA is carried out by a cellular membrane protease with the specificity of furin at the site 164-Arg-Lys-Lys-Arg-167 (Fig. 5) (11). A consensus furin cleavage site has not been found in the Ib sequence.

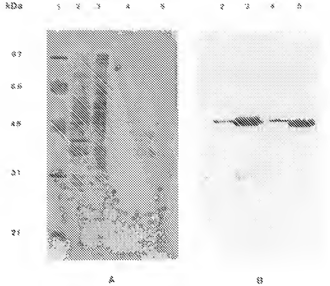

FIG. 6. Analysis of Ia expression in *E. coli* and *C. perfringens*. (A) SDS-PAGE (12% polyacrylamide) of protein extracted from *E. coli* harboring pMRP67 (lane 2) or pMRP91 (lane 3), of culture supernatant of *C. perfringens* (pMRP76) (lane 4), and of purified Ia from recombinant *C. perfringens* (pMRP76) (lane 5). Molecular mass standards are shown in lane 1. (B) Western blot with Ia antibodies. Lanes 2 to 5 are the same as described in panel A.

The hydrophobic sequence Leu-292 to Ser-309 of Ib is predicted to form a transmembrane segment by using the program described by Klein et al. (9). This segment shares a high degree of homology with a sequence found in PA: Leu-284 to Leu-300 (12 identical residues of 17) (Fig. 5). The hydrophobic PA segment is predicted to form an uncertain transmembrane segment (prediction by the quadratic function but not by the linear function of the program of Klein et al. [9]). These data suggest that these sequences could constitute a common functional domain in Ib and PA which could probably be involved in the translocation of the toxin across the cell membrane.

An ATP/GTP binding [consensus GXXXXGK(TS)] (21) has been identified in the Ib sequence from positions Ala-314 to Thr-321 (ASSDQGKT) (Fig. 5). This motif has not been found in the PA sequence.

Expression of Ia gene in recombinant *E. coli* and *C. perfringens*. *E. coli* TG1 harboring the recombinant plasmid pMRP67 and *C. perfringens* S13 harboring pMRP76 (corresponding to the shuttle vector pJIR418 with the pMRP67 insert cut by *Sac*I-*Sal*I localized on the multiple cloning sequence) were analyzed for expression of the Ia gene by Western blotting (immunoblotting) and ADP-ribosylation of actin. As shown in Fig. 6, a 44-kDa peptide recognized by antibodies raised against native purified Ia was found in cell extracts of the recombinant *E. coli* and *C. perfringens* strains and in culture supernatant of *C. perfringens* harboring pMRP76. The affinity-purified Ia from the culture supernatant of the recombinant *C. perfringens* exhibited the same electrophoretic mobility as the native Sa (data not shown).

The pMRP91 corresponding to the subcloning into pACYC184 of the *Eco*RI-*Bcl*I fragment of pMRP67 also induced the production of the Ia component (Fig. 6). These results indicate that the promoter of the Ia gene is localized in the 245 nucleotides upstream of its start codon and that

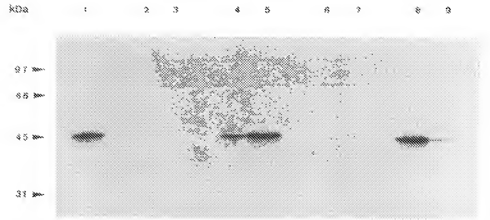

FIG. 7. Autoradiogram of analysis of Ia expression by ADP-ribosylation of actin as described in Materials and Methods. Native Sa as the positive control (lane 1) and cellular extract and culture supernatant of recombinant *C. perfringens* (pMRP76) (lanes 4 and 5, respectively), of recombinant *E. coli* (pMRP67) (lanes 8 and 9, respectively), of *C. perfringens* control (pJIR418) (lanes 2 and 3, respectively), and of *E. coli* control (pUC19) (lanes 6 and 7, respectively) are shown. Molecular mass markers are indicated on the left.

the 1,219 nucleotides upstream of the *Eco*RI site, in pMRP67, are not required for Ia gene expression. Although pMRP67 and pMRP76 contain a part of the Ib gene (DNA fragment encoding the signal peptide, propeptide, and the 99 N-terminal amino acids of the mature Ib), no Ib products, recognized by polyclonal antibodies against Ib, were observed by Western blotting (data not shown).

As shown in Fig. 7, cell extracts and culture supernatants of recombinant *E. coli* and *C. perfringens* strains exhibited ADP-ribosyltransferase activity with actin as the substrate. The enzymatic activity is greater in the cell extract of *E. coli* and the culture supernatant of *C. perfringens* than in the culture supernatant of *E. coli* and the cell extract of *C. perfringens*, indicating that the recombinant Ia is secreted by only *C. perfringens*. The time course of Ia production showed that Ia was released in *C. perfringens* culture supernatant early during the growth phase. In *E. coli*, however, low amounts were released in late growth phase, probably because of spontaneous cell lysis.

Total Ia production (in cell lysate and culture supernatant) was compared in *E. coli* and *C. perfringens* by using a similar concentration of bacteria (10^8 bacteria per ml, counted in a Malassez chamber) in stationary growth phase. Taking into account that the copy number of pJIR418 in *C. perfringens* is 15-fold less than that of pUC19 in *E. coli* (24), it can be estimated that *C. perfringens* produces 8-fold more Ia than *E. coli*.

Expression of the Ib gene in recombinant *E. coli*. The recombinant plasmid pMRP108 was constructed as shown in Fig. 8. This plasmid encompasses the putative promoter upstream from the Ia gene and the Ia and Ib genes. As shown in Fig. 9, the cell extract of *E. coli* TG1 harboring pMRP108 exhibited 90- and 60-kDa peptides recognized by specific anti-Sb antibodies. In length, the 90-kDa peptide corresponded to the recombinant mature Ib and its propeptide. But, the mature Ib is not processed in *E. coli* TG1, indicating that the cleavage site between the propeptide and mature Ib is not recognized by *E. coli* proteases. The Ia gene was also expressed in recombinant *E. coli* TG1 harboring pMRP108 as well as pMRP91 (data not shown).

DISCUSSION

We show that the two independent components Ia and Ib of the iota toxin are encoded by two genes separated by 243 noncoding nucleotides and preceded by a typical Shine-Dalgarno sequence. DNA sequence analysis suggests that the iota-toxin genes are under the control of the same promoter upstream of the Ia gene. A characteristic hydrophobic signal peptide has been identified at the N termini of both Ia and Ib sequences. This is in agreement with the fact that native iota toxin is exported from *C. perfringens* (27).

The iota-toxin sequence seems to be closely related to that of *C. spiroforme* toxin. The protein sequencing of the *C. spiroforme* toxin peptides, which correspond to a total number of 48 residues, shows 83.4% identity with the deduced amino acid sequence of the iota toxin. It has not been excluded that the identity level is higher, since protein sequencing is not as accurate as DNA sequencing. These data are in agreement with the close immunological relationship found between both toxins (16, 19).

The Ia gene has been expressed in *E. coli* and *C. perfringens* under the control of its own promoter. The Ia expression yield was higher in *C. perfringens* than in *E. coli*. The recombinant Ia was recognized by antibodies raised against native Ia and exhibited ADP-ribosyltransferase activity with actin as the substrate. Moreover, the recombinant Ia, in association with native Sb, induced morphological alterations of Vero cells as native iota toxin and *C. spiroforme* toxin did (data not shown). These data indicate that the Ia gene that we have identified corresponds effectively to the Ia structural gene.

The Ib gene was also expressed in *E. coli*, but the processing of the mature Ib did not occur in *E. coli*. It seems that the Ia and Ib genes form an operon under the control of a same promoter upstream from the Ia gene, since no promoter consensus sequence has been found upstream the Ib gene. The apparent lack of expression of the truncated Ib gene in pMRP67 and pMRP76 could be explained by an absence of epitopes recognized by the anti-Sb antibodies. Further work is need to precisely determine the transcription of these two genes.

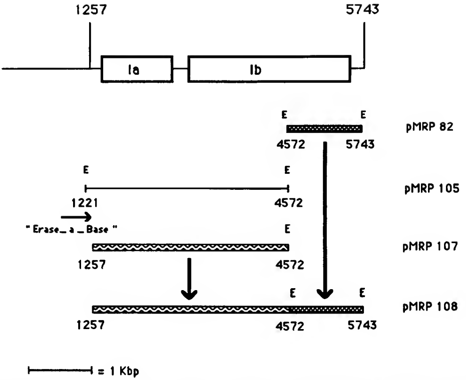

FIG. 8. Construction of recombinant plasmid for Ib expression. The 3,351-bp *Eco*RI fragment encompassing the Ia gene and most of the Ib gene was cloned from *C. perfringens* NCIB 10748 DNA into pUC19 (pMRP105). The *Eco*RI site upstream from the Ia gene was deleted by using the Erase-a-base kit, yielding pMRP107. The *Eco*RI insert of pMRP82 was cloned into the *Eco*RI site of pMRP107 (pMRP108). The numbers indicate the nucleotide positions from Fig. 3.

The Ia sequence shows two residues (Glu-14 and Trp-19) which can be aligned with amino acids involved in the NAD+ binding site of DT, ET, PT, and C3. The consensus motif (Glu/Asp)XXXX-Trp seems to be involved in the NAD+ binding site of these ADP-ribosylating toxins. However, Ia and C3 enzymes do not show a site equivalent to His-21 of DT, which is conserved in the other ADP-ribosylating toxins and has been found to be essential for ADP-ribosylating activity (4). As the substrate of ADP-ribosylation of Ia is an ATP-binding protein (actin) and a GTP-

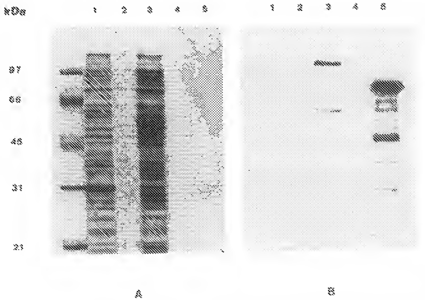

FIG. 9. Analysis of Ib expression in *E. coli*. (A) SDS-PAGE (12% polyacrylamide) of *E. coli* TG1 harboring pUC19 (lanes 1 and 2) or pMRP108 (lanes 3 and 4) (cell extract and culture supernatant, respectively) and of purified Sb (lane 5). (B) Western blot with specific anti-Sb antibodies. The lanes are the same as for panel A.

binding protein for the other ADP-ribosylating toxins, the NAD+ binding site could be slightly different in Ia.

The Ib protein can be divided in three domains (signal peptide, propeptide, and mature protein) defined by two protease cleavage sites. The presence of a propeptide located between the signal peptide and the mature protein is a common feature of most bacterial extracellular proteases, in which it is involved in the correct folding of the mature protein (30). But, no significant homology of Ib with known bacterial protease sequences has been found. The epsilon toxin produced by *C. perfringens* types B and D is secreted as a protoxin (32 kDa), which is trypsin activated by cleavage of 13 basic N-terminal amino acids (7). The propeptide of Ib (172 amino acids) is much longer, and its release does not induce a significant pI change in the protein.

Ib displays a similar structure and significant homology with the PA component of the anthrax toxin (13). Both proteins act as a binding component between cell surface receptors and the enzymatic component and trigger the toxin internalization allowing the delivery of the enzymatic component to its intracellular target. However, PA exhibits some modes of action different from those of Ib. The PA precursor binds to cell surface receptors (the cell recognition domain is located between residues 315 and 735) and is cleaved by a cell surface protease (furin). The propeptide is released, and the mature PA remains bound to the cell surface receptor and is able to bind either the enzymatic component edema factor (an adenylate cyclase) or lethal factor (a metalloprotease) (10, 13). Residues 168 to 312 of PA encompasse the binding site for the edema factor and lethal factor (11, 13). In contrast, Ib needs to be proteolytically activated prior to being functional on cultured cells. These data suggest that the cell surface proteases are probably inactive on Ib. Moreover, the furin cleavage site (Arg-XX-Arg) has not been found in the Ib sequence. The cleavage site Ala-211 to Ala-212 between the propeptide and mature Ib, which has been determined by protein sequencing of the N terminus of mature Ib, corresponds to a common signal peptidase site (29).

The predicted transmembrane segment (Leu-292 to Ser-308) of Ib is highly homologous to the PA segment (Leu-284 to Leu-300) and could be involved in the toxin translocation across the cell membrane. Other binary toxins, such as DT, present hydrophobic helices predicted to be transmembrane segments in their translocation domain (2). In addition, the Ib sequence displays an ATP or GTP binding site (Fig. 5), which could be nonfunctional. However, it is not known whether ATP is required for the entry into the cells of iota toxin.

The relatedness between Ib and PA suggests that they could be derived from a common ancestor gene. Moreover, the GC contents of Ib and PA genes (28 and 31 mol%, respectively) are nearly the same. The central domain implicated in the translocation of the toxin across the cell membrane corresponds to the most conserved region, suggesting that both iota and anthrax toxins probably use a common mechanism of cell entry. The molecular analysis of the other binary toxin genes will provide further information about the evolution of this toxin family.

ACKNOWLEDGMENTS

This work was supported by the Ministère de l'Enseignement Supérieur et de la Recherche and Departement de Microbiologie et Immunologie, Paris XI, UFR Sciences Pharmaceutiques, Châtenay Malabry, France F-92296 (doctoral fellowship to S.P.).

REFERENCES

1. **Chavan, A. J., Y. Nemoto, S. Narumiya, S. Kozaki, and B. E. Haley.** 1992. NAD+ binding site of *Clostridium botulinum* C3 ADP-ribosyltransferase. J. Biol. Chem. **267:**14866–14870.
2. **Choe, S., M. J. Bennett, G. Fujii, P. M. G. Curmi, K. A. Kantardjieff, R. J. Collier, and D. Eisenberg.** 1992. The crystal structure of diphtheria toxin. Nature (London) **357:**216–222.
3. **Considine, R. V., and L. L. Simpson.** 1991. Cellular and molecular actions of binary toxins possessing ADP-ribosyltransferase activity. Toxicon **29:**913–936.
4. **Domenighini, M., C. Montecucco, W. C. Ripka, and R. Rappuoli.** 1991. Computer modelling of the NAD binding site of ADP-ribosylating toxins; active-site structure and mechanism of NAD binding. Mol. Microbiol. **5:**23–31.
5. **Guyonnet, F., and M. Plommet.** 1970. Hemolysine gamma de *Staphylococcus aureus*: purification et propriétés. Ann. Inst. Pasteur **118:**19–33.
6. **Hatheway, C. L.** 1990. Toxigenic clostridia. Clin. Microbiol. Rev. **3:**66–98.
7. **Hunter, S. E. C., I. N. Clarke, D. C. Kelly, and R. W. Titball.** 1992. Cloning and nucleotide sequencing of the *Clostridium perfringens* epsilon-toxin gene and its expression in *Escherichia coli*. Infect. Immun. **60:**102–110.
8. **Inoue, S., M. Sugai, Y. Murooka, S. Y. Paik, Y. M. Hong, H. Ohgai, and H. Suginaka.** 1991. Molecular cloning and sequencing of the epidermal cell differentiation inhibitor gene from *Staphylococcus aureus*. Biochem. Biophys. Res. Commun. **174:**459–464.
9. **Klein, P., M. Kanehisha, and C. DeLisl.** 1985. The detection and classification of membrane-spanning proteins. Biochim. Biophys. Acta **815:**468–476.
10. **Klimpel, K. R., N. Arora, and S. Leppla.** 1993. Anthrax toxin lethal factor has homology to the thermolysin-like proteases and displays proteolytic activity, abstr. B-111, p. 45. Abstr. 93rd Gen. Meet. Am. Soc. Microbiol. 1993. American Society for Microbiology, Washington, D.C.
11. **Klimpel, K. R., S. S. Molloy, G. Thomas, and S. H. Leppla.** 1992. Anthrax toxin protective antigen is activated by a cell surface protease with the sequence specificity and catalytic properties of furin. Proc. Natl. Acad. Sci. USA **89:**10277–10281.
12. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
13. **Leppla, S.** 1991. The anthrax toxin complex, p. 277–302. *In* J. E. Alouf and J. Freer (ed.), Sourcebook of bacterial protein toxins. Academic Press, New York.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Noda, M., T. Mirayama, I. Kato, and F. Matsuda.** 1980. Crystallisation and properties of staphylococcal leucocidin. Biochim. Biophys. Acta **633:**33–39.
16. **Popoff, M. R., and P. Boquet.** 1988. *Clostridium spiroforme* toxin is a binary toxin which ADP-ribosylates cellular actin. Biochem. Biophys. Res. Commun. **152:**1361–1368.
17. **Popoff, M. R., D. Hauser, P. Boquet, M. W. Eklund, and D. M. Gill.** 1991. Characterization of the C3 gene of *Clostridium botulinum* types C and D and its expression in *Escherichia coli*. Infect. Immun. **59:**3673–3679.
18. **Popoff, M. R., E. J. Rubin, D. M. Gill, and P. Boquet.** 1988. Actin-specific ADP-ribosyltransferase produced by a *Clostridium difficile* strain. Infect. Immun. **56:**2299–2306.
19. **Popoff, M. R., F. W. Milward, B. Bancillon, and P. Boquet.** 1989. Purification of the *Clostridium spiroforme* binary toxin and activity of the toxin on HEp-2 cells. Infect. Immun. **57:**2462–2469.
20. **Pugsley, A. P.** 1989. Protein targeting. Academic Press, San Diego, Calif.
21. **Saraste, M., P. R. Sibbald, and A. Wittinghofer.** 1990. The P-loop-a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. **15:**430–434.
22. **Scott, P. T., and J. I. Rood.** 1989. Electroporation-mediated transformation of lysostaphin-treated *Clostridium perfringens*.

Gene **82:**327–333.

23. **Simpson, L. L., B. G. Stiles, H. H. Zepeda, and T. D. Wilkins.** 1987. Molecular basis for the pathological actions of *Clostridium perfringens* iota toxin. Infect. Immun. **55:**118–122.
24. **Sloan, J., T. A. Warner, P. T. Scott, T. L. Bannam, D. I. Berryman, and J. I. Rood.** 1992. Construction of a sequenced *Clostridium perfringens-Escherichia coli* shuttle plasmid. Plasmid **27:**207–219.
25. **Spangler, B. D.** 1992. Structure and function of cholera toxin and the related *Escherichia coli* heat-labile enterotoxin. Microbiol. Rev. **56:**622–647.
26. **Stiles, B. G., and T. D. Wilkins.** 1986. Purification and characterization of *Clostridium perfringens* iota toxin: dependence on two nonlinked proteins for biological activity. Infect. Immun. **54:**683–688.
27. **Stiles, B. G., and T. D. Wilkins.** 1986. *Clostridium perfringens* iota toxin: synergism between two proteins. Toxicon **24:**767–773.
28. **Vandekerckhove, J., B. Schering, M. Bärmann, and K. Aktories.** 1987. *Clostridium perfringens* iota toxin ADP-ribosylates skeletal muscle actin in Arg-177. FEBS Lett. **225:**48–52.
29. **VonHeijne, G., and L. Abrahmsen.** 1989. Species-specific protein secretion in signal peptide design. Implication for protein secretion in foreign hosts. FEBS Lett. **244:**439–446.
30. **Wandersman, C.** 1989. Secretion, processing and activation of bacterial extracellular proteases. Mol. Microbiol. **3:**1825–1831.
31. **Young, M., N. P. Minton, and W. L. Staudenbauer.** 1989. Recent advances in the genetics of the clostridia. FEMS Microbiol. Lett. **63:**301–326.